IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Weissman   Art Unit   : 2166
Serial No. : 10/748,505   Examiner : Sangwoo Ahn
Filed   : December 30, 2003 Conf. No. : 7264
Title   : METHODS AND SYSTEMS FOR COMPRESSING INDICES

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

BRIEF ON APPEAL

This Brief on Appeal perfects the Notice of Appeal filed herewith and appeals the

rejections set forth in the Office action mailed June 25, 2007.

(1)   Real Party in Interest

This case is assigned of record to Google Inc., who is the real party in interest.

(2)   Related Appeals and Interferences

There are no known related appeals and/or interferences.

(3)   Status of Claims

Claims 1-35 are pending.

Claims 1-35 are under consideration.

No claims have been canceled.

Claims 1-35 stand rejected.

Claim 1, 11, 16, 18, and 28 are in independent form.

Claims 1, 11, 16, 18, 28, 30, 31, 33, and 34 are involved directly in the appeal. Claims 2-

10, 12-15, 17, 19-27, 29, 32, and 35 are not directly involved in the appeal but rather are

involved only by virtue of their dependency from one or more of claims 1, 11, 16, 18, 28, 30, 31,

33, and 34.

(4) Status of Amendments

All claim amendments have been entered.

(5) Summary of Claimed Subject Matter

Claim 1 relates to a method that includes:

selecting from an inverted index (*See, e.g., specification*, page 7, line 8-10; page 8, line 10-page 9, line 1) at least a first item entry comprising a first listing of articles that are associated with a first item and a second item entry comprising a second listing of articles that are associated with a second item (*See id. See also, e.g., specification*, page 12, line 17-19; page 13, line 13-15), wherein the second item differs from the first item (*See, e.g., specification*, page 12, line 8-10; page 13, line 4-8, line 16-17; page 13, line 13-15);

determining whether to compress the second item entry into the first item entry (*See, e.g., specification*, page 12, line 21-page 13, line 10; page 13, line 16-page 15, line 15); and

compressing the second item entry into the first item entry based on the determination. *See, e.g., specification*, page 15, line 16-page 16, line 22.

Claim 11 relates to a method that includes:

selecting from an inverted index (*see, e.g., specification*, page 7, line 8-10; page 8, line 10-page 9, line 1) at least a first item entry comprising a first listing of articles that are associated with a first item (*see id. See also, e.g., specification*, page 12, line 17-19; page 13, line 13-15) and an item value for each article in the first listing (*See, e.g., specification*, page 9, line 22- page 10, line 3), and

a second item entry comprising a second listing of articles that are associated with a second item (*see id. See also, e.g., specification*, page 12, line 17-19; page 13, line 13-15) and

an item value for each article in the second listing (*See, e.g., specification*, page 9, line 22- page

10, line 3), wherein the second item differs from the first item (*See, e.g., specification*, page 12,

line 8-10; page 13, line 4-8, line 16-17; page 13, line 13-15);

determining a cost-benefit ratio for compressing the second item entry into the first item

entry (*see, e.g., specification*, page 12, line 21-page 13, line 10; page 13, line 16-page 15, line

15);

comparing the cost-benefit ratio with a value to determine if the cost-benefit ratio is

acceptable (*See, e.g., specification*, page 13, line 1-4; page 16, line 5-13); and

if the cost-benefit ratio is acceptable, compressing the second item entry into the first

item entry. S*ee, e.g., specification*, page 15, line 16-page 16, line 22.

Claim 16 relates to a method that includes:

selecting from an inverted index (*See, e.g., specification*, page 7, line 8-10; page 8, line

10-page 9, line 1) a plurality of item entries, each item entry different from the other selected

item entries (*See, e.g., specification*, page 12, line 8-10; page 13, line 4-8, line 16-17; page 13,

line 13-15) and each item entry comprising a listing of associated articles (*See specification*,

page 7, line 8-10; page 8, line 10-page 9, line 1 *See also, e.g., specification*, page 12, line 17-19;

page 13, line 13-15);

determining whether to compress the plurality of item entries (*See, e.g., specification*,

page 12, line 21-page 13, line 10; page 13, line 16-page 15, line 15); and

compressing the item entries based on the determination. *See, e.g., specification*, page

15, line 16-page 16, line 22.

Claim 18 relates to an article comprising one or more computer-readable media containing program code operable to cause one or more machines to perform operations (*See, e.g., specification*, page 4, line 15-page 5, line 3; page 6, line 3-5, 12-15), the operations comprising:

selecting from an inverted index (*See, e.g., specification*, page 7, line 8-10; page 8, line 10-page 9, line 1) at least a first item entry comprising a first listing of articles that are associated with a first item, and a second item entry comprising a second listing of articles that are associated with a second item (*See specification*, page 7, line 8-10; page 8, line 10-page 9, line 1 *See also, e.g., specification*, page 12, line 17-19; page 13, line 13-15), wherein the second item differs from the first item (*See, e.g., specification*, page 12, line 8-10; page 13, line 4-8, line 16-17; page 13, line 13-15);

determining whether to compress the second item entry into the first item entry (*See, e.g., specification*, page 12, line 21-page 13, line 10; page 13, line 16-page 15, line 15); and

compressing the second item entry into the first item entry based on the determination. *See, e.g., specification*, page 15, line 16-page 16, line 22.

Claim 28 relates to an article comprising one or more computer-readable media containing program code operable to cause one or more machines to perform operations (*See, e.g., specification*, page 4, line 15-page 5, line 3; page 6, line 3-5, 12-15), the operations comprising:

selecting from an inverted index a plurality of item entries (*see, e.g., specification*, page 7, line 8-10; page 8, line 10-page 9, line 1), each item entry different from the other selected item entries (*See, e.g., specification*, page 12, line 8-10; page 13, line 4-8, line 16-17; page 13, line 13-

15) and each item entry comprising a listing of associated articles (*See specification*, page 7, line 8-10; page 8, line 10-page 9, line 1 *See also, e.g., specification*, page 12, line 17-19; page 13, line 13-15);

determining whether to compress the plurality of item entries (*See, e.g., specification*, page 12, line 21-page 13, line 10; page 13, line 16-page 15, line 15); and

compressing the item entries based on the determination. *See, e.g., specification*, page 15, line 16-page 16, line 22.

Claim 30 relates to the method of claim 1, wherein:

the first item comprises a first word (*See, e.g., specification*, page 1, line 24);

the articles in the first listing are associated with the first item by virtue of the first word appearing in the articles in the first listing (*See, e.g., specification*, page 1, line 25-page 2, line 1. *See also* page 2, line 4-5, line 7-10); and

the first word does not appear in the second item. *See, e.g., specification*, page 12, line 8-10; page 13, line 4-8, line 16-17; page 13, line 13-15.

Claim 31 relates to the method of claim 1, wherein:

the first item comprises a first concept (*See, e.g., specification*, page 1, line 24; page 2, line 4-10; page 8, line 10-page 9, line 1);

the second item comprises a second concept (*See, e.g., specification*, page 1, line 24; page 2, line 4-10; page 8, line 10-page 9, line 1);

the articles in the first listing are associated with the first item by virtue of the first concept appearing in the articles in the first listing (*See, e.g., specification*, page 1, line 25-page

2, line 1. *See also* page 2, line 4-5, line 7-10, page 8, line 10-page 9, line 1);

the articles in the second listing are associated with the second item by virtue of the second concept appearing in the articles in the second listing (*See, e.g., specification*, page 1, line 25-page 2, line 1. *See also* page 2, line 4-5, line 7-10, page 8, line 10-page 9, line 1); and

determining whether to compress the second item entry into the first item entry comprises determining whether the first concept is related to the second concept. *See, e.g., specification*, page 12, line 10-17.

Claim 33 relates to the article of claim 18, wherein:

the first item comprises a first word (*See, e.g., specification*, page 1, line 24);

the articles in the first listing are associated with the first item by virtue of the first word appearing in the articles in the first listing (*See, e.g., specification*, page 1, line 25-page 2, line 1. *See also* page 2, line 4-5, line 7-10); and

the first word does not appear in the second item. *See, e.g., specification*, page 12, line 8-10; page 13, line 4-8, line 16-17; page 13, line 13-15.

Claim 34 relates to the article of claim 18, wherein:

the first item comprises a first concept (*See, e.g., specification*, page 1, line 24; page 2, line 4-10; page 8, line 10-page 9, line 1);

the second item comprises a second concept (*See, e.g., specification*, page 1, line 24; page 2, line 4-10; page 8, line 10-page 9, line 1);

the articles in the first listing are associated with the first item by virtue of the first concept appearing in the articles in the first listing (*See, e.g., specification*, page 1, line 25-page

2, line 1. *See also* page 2, line 4-5, line 7-10, page 8, line 10-page 9, line 1);

the articles in the second listing are associated with the second item by virtue of the second concept appearing in the articles in the second listing (*See, e.g., specification*, page 1, line 25-page 2, line 1. *See also* page 2, line 4-5, line 7-10, page 8, line 10-page 9, line 1); and

determining whether to compress the second item entry into the first item entry comprises determining whether the first concept is related to the second concept. *See, e.g., specification*, page 12, line 10-17.

(6) Grounds of Rejection to be Reviewed on Appeal

As set forth in the following concise statements, the following grounds for rejection are presented for review on appeal:

Ground 1: Whether claims 1, 9, 16-18, 26, 28-31, and 33-34[1] are properly rejected under 35 U.S.C. § 102(e) as anticipated by U.S. Patent Publication No. 2003/0088715 to Chaudhuri et al. (hereinafter "Chaudhuri").

Ground 2: Whether claims 3-7, 11-15, and 20-24 are properly rejected under 35 U.S.C. § 103(a) as obvious over Chaudhuri, U.S. Patent No. 6,834,290 to Pugh et al. (hereinafter "Pugh"), and U.S. Patent No. 5,915,249 to Spencer (hereinafter "Spencer").

---

[1] Please note that although the boldfaced heading on page 5 of the Office action mailed June 25, 2007 does not identify that claims 30-31 and 33-34 have been rejected under 35 U.S.C. § 102(e) as anticipated by Chaudhuri, page 6-7 does describe such a rejection. Accordingly, Applicant has assumed that the boldfaced heading on page 5 is incorrect. If this assumption is incorrect, Applicant respectfully requests that a new action on the merits be issued to clarify the status of claims 30-31 and 33-34 .

(7)   **Argument**

The organization of the arguments presented hereinafter follows the organization of the grounds for rejection to be reviewed on appeal set forth above. In particular, a separate boldfaced heading for each ground presented for review follows.

Ground 1: Rejections under 35 U.S.C. § 102(e)

Claims 1 and 18 were rejected under 35 U.S.C. § 102(e) as anticipated by Chaudhuri.

As discussed above, claim 1 relates to a method that relates to the compression of item entries in an inverted index. The recited item entries each comprise a listing of articles associated with an item. Moreover, the items with which the articles are associated differ. Claim 18 relates to an article that includes one or more computer-readable media containing program code operable to cause one or more machines to perform related operations.

Chaudhuri's compression of a symbol table neither describes nor suggests such a compression of item entries. Indeed, Chaudhuri's compression is understood to _yield_ a symbol table in which item entries are associated with multiple locations.

In this regard, Chaudhuri deals with the searching of a relational database using an index of the data records in the relational database. *See, e.g., Chaudhuri*, para. [0008]. Such an index maps a data record to a region of the database in which the record is found. *Id.*, para. [0009]. For keyword searching, a symbol table is constructed that stores location information for each keyword. *Id.*, para. [0009].

Chaudhuri's FIG. 4 is an example of such a symbol table, and FIG. 5 is a compressed version of this symbol table. *Id.*, paras. [0015], [0016]. For the sake of convenience, these FIGS. are now reproduced.

133

HashVal	ColId
v_1	c_1
v_2	c_1
v_3	c_1
v_4	c_1
v_5	c_2
v_3	c_2
v_4	c_2
v_5	c_2

Fig.4

134

HashVal	ColId
v_1	c_1
v_2	x
v_3	x
v_4	x
v_5	c_2

135

NewColId	ColId
x	c_1
x	c_2

Fig.5

Rather than store keywords directly in these symbol tables, the keywords are hashed into hash

values (i.e., HashVal's) that are stored in the symbol tables, along with identifiers of the columns

(i.e., ColId's) in which the keywords occurred. *Id.*, para. [0034]. These symbol tables thus

associate "keywords and the corresponding database tables and columns in which the keyword

can be found." *Id.*, para. [0031].

As can be seen from FIG. 4, some keywords appear at multiple locations in the relational

database. For example, the keyword associated with the hash value "v_2" appears in the columns

"c_1" and "c_2" in the relational database. The hash value "v_2" has multiple entries in the symbol

table of FIG. 4, one for each of these locations. Chaudhuri describes that the multiple entries for

such keywords can be compressed to yield the symbol table shown in FIG. 5. *Id.*, para. [0035].

In the compressed symbol table 134, the ColId "x" represents the conjunction of "c_1" and "c_2," as

enumerated in ColumnsMap table 135. *Id.*

The rejection is based on the contention that this compression constitutes a compression

of one item entry into another item entry. However, claims 1 and 18 also recite that:

-the first item entry comprises a first listing of articles associated with a first item;

-the second item entry comprises a second listing of articles associated with a second item; and

-the second item differs from the first item.

None of these elements and/or limitations is described or suggested by Chaudhuri's compression. As discussed, Chaudhuri's compression compresses associations of single hash values with single ColId's to associate single hash values with multiple ColId's. Before compression, none of the hash values has a "listing of articles" associated with them. Instead, each hash value has a *single* ColId associated with it. Chaudhuri's compression thus must occur before any selection of item entries that comprise a first listing of articles from an inverted index, as recited in claims 1 and 18. Chaudhuri thus neither describes nor suggests selecting first and second items from an inverted index, as recited in claims 1 and 18.

Moreover, it is clear that the single hash values which are compressed in Chaudhuri are identical to one another. For example, the hash value "v_2" that is associated with "c_1" is identical to the hash value "v_2" that is associated with "c_2." Since these hash values are identical to one another, Chaudhuri neither describes nor suggests compressing a second item entry into a first item entry where the second item differs from the first item, as recited in claims 1 and 18.

After these deficiencies were pointed out in the response to Office action filed February 20, 2007 and during the telephone interview on February 20, 2007, the Examiner extended the courtesy of spelling out the basis of the rejection in tabular form. *See Office action mailed June 25, 2007*, page 3-4. Applicant thanks the Examiner for the courtesy. Nevertheless, Applicant respectfully disagrees with the basis of the rejection set forth in the table for several reasons.

For example, the first row of the table contends that Chadhuri's FIG. 4 shows column

ID's C1 and C2 as constituting a "listing of articles" for hash value V2. Applicant respectfully

disagrees. In this regard, claims 1 and 18 recite that first and second item entries in an inverted

index <u>each</u> comprise a listing of articles. However, in FIG. 4, column ID's C1 and C2 are plainly

shown in association with hash value V2 in different entries (i.e., the second entry and the fifth

entry) in Chadhuri's symbol table. This interpretation of the term "entry" in the context of an

inverted index is believed to be well understood to those of ordinary skill. *See, e.g., U.S. Patent*

No. 5,915,249 to Spencer (cited in rejecting claim 11 and discussed further below). Thus, in

Chadhuri's FIG. 4, each entry includes a hash value that is associated with a *single* column ID

rather than to a listing of articles, as recited in claims 1 and 18. Indeed, as discussed above, the

objective of Chadhuri's compression is to arrive at a symbol table in which a hash value is

associated with multiple column ID's, via the NewColId's shown in FIG. 5.

As another example, the final row of the table contends that "recurring column ID's are

compressed…" This is presumably based upon Chadhuri's description that entries for hash

values that are "in common" are compressed. *See Chadhuri*, para. [0035]. In other words,

Chadhuri's technique compresses entries the same "hash values that occur in [column] c1 as well

as [column] c2." *Id.*

However, claims 1 and 18 recite that a second item entry is compressed into a first item

entry, where the second item differs from the first item. Chadhuri's compression of multiple

entries that have a common hash value into a single entry neither describes nor suggests the

recited compression. In particular, entries with hash values that are "in common" (e.g., different

entries with the identical hash value "v2") are not entries where a second item differs from a first item. Instead, the hash values "in common" are understood to be identical to one another.

Since elements and/or limitations recited in claims 1 and 18 are neither described nor suggested by Chaudhuri, claims 1 and 18 are not anticipated by Chaudhuri. Accordingly, Applicant requests that the rejections of claims 1, 18, and the claims dependent therefrom be withdrawn.

Claims 16 and 28 were rejected under 35 U.S.C. § 102(e) as anticipated by Chaudhuri.

As discussed above, claim 16 relates to a method that includes selecting and compressing a plurality of item entries from an inverted index, wherein each item entry is different from the other selected item entries and each item entry comprises a listing of associated articles. Claim 28 relates to an article that includes one or more computer-readable media containing program code operable to cause one or more machines to perform related operations.

The rejection of claims 16 and 28 do not set forth a separate basis for the rejection of claims 16 and 28 but rather states that claims 16 and 28 are rejected "on the same rationale" used in rejecting claim 1. Applicant respectfully disagrees with the rejections.

In this regard, as discussed above, Chaudhuri describes that entries in a symbol table that each have the same hash value and are each associated with a single ColId can be compressed to yield a combined entry for that hash value that is associated with multiple ColId's.

Since each of Chaudhuri's hash values initially has a single ColId associated with it, Chaudhuri neither describes nor suggests selecting and compressing a plurality of item entries from an inverted index, wherein each item entry is different from the other selected item entries and each item entry comprises a listing of associated articles, as recited in claims 16 and 28.

Indeed, the objective of Chaudhuri's compression would appear to be the generation of a symbol table in which a hash value is associated with multiple column ID's (via the NewColId's shown in FIG. 5).

Moreover, since the single hash values which are compressed in Chaudhuri are "in common," Chaudhuri neither describes nor suggests compressing the item entries where each item entry is different from other selected item entries, as recited in claims 16 and 28.

Since elements and/or limitations recited in claims 16 and 28 are neither described nor suggested by Chaudhuri, claims 16 and 28 are not anticipated by Chaudhuri. Accordingly, Applicant requests that the rejections of claims 16, 28, and the claims dependent therefrom be withdrawn.

Claims 30 and 33 were rejected under 35 U.S.C. § 102(e) as anticipated by Chaudhuri.

As discussed above, claims 30 and 33 recite that the first item comprises a first word and the first word does not appear in the second item.

The rejection of claims 30 and 33 point out that each of Chaudhuri's hash values represents a different keyword. *See Office action mailed June 25, 2007*, page 6. While this is true, applicant would like to point out that Chaudhuri's entries with hash values that represent different keywords *are not compressed into one another*. Instead, as discussed above, only entries that have identical hash values (i.e., entries with hash values "in common") are compressed into each other. Since, by virtue of their dependency from their respective independent claims, claims 30 and 33 both recite that a second item entry is compressed into a first item entry. Thus, claims 30 and 33 are not anticipated by Chaudhuri on this basis as well.

Since elements and/or limitations recited in claims 30 and 33 are neither described nor suggested by Chaudhuri, claims 30 and 33 are not anticipated by Chaudhuri. Accordingly, Applicant requests that the rejections of claims 30 and 33 be withdrawn.

Claims 31 and 34 were rejected under 35 U.S.C. § 102(e) as anticipated by Chaudhuri.

As discussed above, claims 31 and 34 recite that a first item comprises a first concept and a second items comprises a second concept. By virtue of their dependency from their respective independent claims, claims 31 and 34 both recite that the first item differs from the second item.

As best understood, the rejection of claims 31 and 34 contends that the different keywords represented by each of Chaudhuri's hash values constitute first and second concepts, as recited. *See Office action mailed June 25, 2007*, page 6.

Applicant respectfully disagrees. Chaudhuri's entries with hash values that represent different keywords *are not compressed into one another*. Instead, as discussed above, only entries that have identical hash values (i.e., entries with hash values "in common") are compressed into each other. Since, by virtue of their dependency from their respective independent claims, claims 31 and 34 both recite that a second item entry is compressed into a first item entry. Thus, claims 31 and 34 are not anticipated by Chaudhuri on this basis as well.

Since elements and/or limitations recited in claims 31 and 34 are neither described nor suggested by Chaudhuri, claims 31 and 34 are not anticipated by Chaudhuri. Accordingly, Applicant requests that the rejections of claims 31, 34, and the claims dependent therefrom be withdrawn.

Ground 2: Rejections under 35 U.S.C. § 103(a)

Claim 11 was rejected under 35 U.S.C. § 103(a) as obvious over Chaudhuri, Pugh, and

Spencer.

As discussed above, claim 11 relates to a method that includes selecting and compressing

item entries, from an inverted index, that comprise listings of articles and item values for each

article in the listings. Claim 11 also recites that a cost-benefit ratio for compressing a second

item entry into a first item entry is determined.

The rejection of claim 11 is based on the contention that it would have been obvious for

one of ordinary skill to have combined Chaudhuri, Pugh, and Spencer and arrive at the recited

subject matter. Applicant respectfully disagrees for several reasons.

For example, none of Chaudhuri, Pugh, and Spencer compress item entries that comprise

listings of articles and item values for each article in the listings, as recited. In this regard, as

discussed above, Chaudhuri describes that entries in a symbol table that each have the same hash

value and are each associated with a single ColId can be compressed to yield a combined entry

for that hash value that is associated with multiple ColId's. Since each of Chaudhuri's hash

values initially has a single ColId associated with it, Chaudhuri neither describes nor suggests

selecting and compressing a plurality of item entries from an inverted index, wherein each item

entry is different from the other selected item entries and each item entry comprises a listing of

associated articles, as recited in claim 11. Indeed, the objective of Chaudhuri's compression

would appear to be the generation of a symbol table in which a hash value is associated with

multiple column ID's (via the NewColId's shown in FIG. 5). Moreover, since the entries in

Chaudhuri which are compressed have identical hash values (i.e., hash values "in common"),

Chaudhuri neither describes nor suggests compressing the item entries where each item entry is different from other selected item entries, as recited in claim 11.

Pugh and Spencer do nothing to remedy these deficiencies in Chaudhuri. As for Pugh, Pugh's primary concern is with the reorganization of data in a database. *See, e.g., Pugh*, col. 1, line 21-23. According to Pugh, the reorganizing a database "will generally eliminate wasted blocks in the [database] object, reduce the number of chained rows, and reduce the number of extents allocated. The free space distribution in a tablespace is improved by, for example, causing free space to be organized into fewer, larger extents as opposed to many, smaller extents." *See, e.g., Pugh*, col. 10, line 43-51.

Applicant respectfully submits that Pugh's reorganization of databases will not make the selection and compression of item entries recited in claim 1 obvious to those of ordinary skill. In this regard, database reorganization is understood to reorganize a database rather than compress item entries, as recited. Indeed, to the extent that Pugh can be read as advocating data reorganization, Applicant submits that Pugh would lead those of ordinary skill away from the recited subject matter. In this regard, since Pugh describes how to reorganize databases, applicant is at a loss to understand why those of ordinary skill would turn to compression of item entries in an inverted index.

As for Spencer, Spencer's primary concern is with efficient information retrieval from a database. *See, e.g., Spencer*, col. 1, line 46-49. Spencer describes that a static cache of "important" documents can be used to supplement searches of an inverted index. *See, e.g., id.*, col. 1, line 49-53. *See also id.* col. 3, line 33-52. The static cache can be a primary index of the inverted index. *See, e.g., id.*, col. 4, line 20-21. The static cache contains an entry for each term

of the inverted index that exceeds a certain number of document tuples. *See, e.g., id.*, col. 3, line

36-38. The static cache thus does not compress entries in the inverted index, but rather includes

only a subset of the entries in the inverted index.

Applicant respectfully submits that Spencer's static cache will not make the selection and

compression of item entries recited in claim 1 obvious to those of ordinary skill. In this regard,

the entries in Spencer's static cache are understood to correspond to a subset of entries in a larger

inverted index. The entries in the larger inverted index are not compressed, as recited, but rather

merely reduced in number. Indeed, to the extent that Spencer can be read as advocating the use

of a static cache with a reduced number of entries, Applicant submits that Spencer would lead

those of ordinary skill away from the recited subject matter. In this regard, since Spencer

describes that a static cache with a reduced number of entries can be used to speed information

retrieval, applicant is at a loss to understand why those of ordinary skill would turn to

compression of item entries in an inverted index.

Accordingly, even if Chaudhuri, Pugh, and Spencer were combined, one of ordinary skill

would not arrive at the recited selection and compression of item entries that comprise listings of

articles and item values for each article in the listings.

In addition, Applicant submits that none of the cited references would lead one of

ordinary skill to determine a cost-benefit ratio for compressing item entries that comprise listings

of articles and item values for each article in the listings. In this regard, Chaudhuri only

describes the compression of entries where the hash values are identical. Chaudhuri does not

describe that there is any cost associated with such a compression, much less how to compute

such a cost. Pugh and Spencer have nothing to do with the compression of entries in an inverted

index and also do not recognize any cost associated with such a compression. Accordingly, even if Chaudhuri, Pugh, and Spencer were combined, one of ordinary skill would not arrive at the recited determination of a cost-benefit ratio for compressing a second item entry into a first item entry.

Since one of ordinary skill would not arrive at the subject matter recited in claim 11 even if Chaudhuri, Pugh, and Spencer were combined, claim 11 is not obvious over Chaudhuri, Pugh, and Spencer in any combination. Accordingly, Applicant requests that the rejections of claim 11 and the claims dependent therefrom be withdrawn.

It is believed that all of the pending claims have been addressed. However, the absence of a reply to a specific rejection, issue, or comment does not signify agreement with or concession of that rejection, issue, or comment. In addition, because the arguments made above may not be exhaustive, there may be reasons for patentability of any or all pending claims (or other claims) that have not been expressed. Finally, nothing in this paper should be construed as an intent to concede any issue with regard to any claim, except as specifically stated in this paper, and the amendment of any claim does not necessarily signify concession of unpatentability of the claim prior to its amendment.

The brief fee of $510 is enclosed. Please apply any other charges or credits to Deposit

Account No. 06-1050.

Respectfully submitted,

Date:10/25/2007   /John F. Conroy, Reg. # 45,485/
                                        John F. Conroy
                                        Reg. No. 45,485

Fish & Richardson P.C.
12390 El Camino Real
San Diego, California 92130
Telephone: (858) 678-5070
Facsimile: (858) 678-5099

Appendix of Claims

1. A method, comprising:

selecting from an inverted index at least

   a first item entry comprising a first listing of articles that are associated with a first item and

   a second item entry comprising a second listing of articles that are associated with a second item, wherein the second item differs from the first item;

determining whether to compress the second item entry into the first item entry; and

compressing the second item entry into the first item entry based on the determination.

2. The method of claim 1, wherein determining whether to compress the second item entry into the first item entry comprises:

determining a cost-benefit ratio for compressing the second item entry into the first item entry; and

comparing the cost-benefit ratio with a value to determine if the cost-benefit ratio is acceptable.

3. The method of claim 1, wherein:

the first item entry further comprises an item value for each article in the first listing; and

the second item entry further comprises an item value for each article in the second listing.

4. The method of claim 3, wherein the item values comprise representations of strengths of the items in the articles.

5. The method of claim 3, wherein the item values comprise representations of whether the items appear in the articles.

6. The method of claim 2, wherein a cost in the cost-benefit ratio comprises a representation of:

a loss in precision caused by compressing the first item entry and the second item entry; or

additional processing time required when utilizing a compressed entry.

7. The method of claim 2, wherein determining the cost-benefit ratio comprises determining how much the first item entry and the second item entry change with the second item entry compressed into the first item entry.

8. The method of claim 2, wherein a benefit in the cost-benefit ratio comprises a representation of an amount of memory saved by compression of the first item entry and the second item entry.

9. The method of claim 1, wherein the items comprise one or more of words, concepts, or images.

10. (Currently Amended) The method of claim 2, wherein the value is predetermined.

11. A method, comprising:

selecting from an inverted index at least

a first item entry comprising a first listing of articles that are associated with a first item and an item value for each article in the first listing, and

a second item entry comprising a second listing of articles that are associated with a second item and an item value for each article in the second listing, wherein the second item differs from the first item;

determining a cost-benefit ratio for compressing the second item entry into the first item entry;

comparing the cost-benefit ratio with a value to determine if the cost-benefit ratio is acceptable; and

if the cost-benefit ratio is acceptable, compressing the second item entry into the first item entry.

12. The method of claim 11, wherein determining the cost-benefit ratio comprises determining how much the first item entry and the second item entry change with the second item entry compressed into the first item entry.

13. The method of claim 11, wherein a benefit in the cost-benefit ratio comprises a representation of an amount of memory saved by compression of the first item entry and the second item entry.

14. The method of claim 11, wherein the value is predetermined.

15. The method of claim 11, wherein the items comprise one or more of words, concepts, or images.

16. A method, comprising:

selecting from an inverted index a plurality of item entries, each item entry different from the other selected item entries and each item entry comprising a listing of associated articles;

determining whether to compress the plurality of item entries; and

compressing the item entries based on the determination.

17. The method of claim 16, wherein the plurality of item entries comprises three or more item entries.

18. An article comprising one or more computer-readable media containing program code operable to cause one or more machines to perform operations, the operations comprising:

selecting from an inverted index at least

a first item entry comprising a first listing of articles that are associated with a first item, and

a second item entry comprising a second listing of articles that are associated with a second item, wherein the second item differs from the first item;

determining whether to compress the second item entry into the first item entry; and

compressing the second item entry into the first item entry based on the determination.

19. The article of claim 18, wherein determining whether to compress the second item entry into the first item entry comprises:

determining a cost-benefit ratio for compressing the second item entry into the first item

entry; and

comparing the cost-benefit ratio with a value to determine if the cost-benefit ratio is

acceptable.

20. The article of claim 18, wherein:

the first item entry further comprises an item value for each article in the first listing; and

the second item entry further comprises an item value for each article in the second

listing.

21. The article of claim 20, wherein the item values comprise representations of

strengths of the items in the articles.

22. The article of claim 20, herein the item values comprise representations of whether

the items appear in the articles.

23. The article of claim 19, wherein a cost in the cost-benefit ratio comprises a

representation of:

a loss in precision caused by compressing the first item entry and the second entry; or

additional processing time required when utilizing a compressed entry.

24. The article of claim 19, wherein determining the cost-benefit ratio comprises

determining how much a first item entry and a second item entry change with the second item

entry compressed into the first item entry.

25. The article of claim 19, wherein a benefit for the cost-benefit ratio comprises a representation of an amount of memory saved by compression of the first item entry and the second item entry.

26. The article of claim 18, wherein the items comprise one or more of words, concepts, or images.

27. The article of claim 19, wherein the value is predetermined.

28. An article comprising one or more computer-readable media containing program code operable to cause one or more machines to perform operations, the operations comprising:

selecting from an inverted index a plurality of item entries, each item entry different from the other selected item entries and each item entry comprising a listing of associated articles;

determining whether to compress the plurality of item entries; and

compressing the item entries based on the determination.

29. The article of claim 28, wherein the plurality of item entries comprises three or more item entries.

30. The method of claim 1, wherein:

the first item comprises a first word;

the articles in the first listing are associated with the first item by virtue of the first word appearing in the articles in the first listing; and

the first word does not appear in the second item.

31. The method of claim 1, wherein:

the first item comprises a first concept;

the second item comprises a second concept;

the articles in the first listing are associated with the first item by virtue of the first

concept appearing in the articles in the first listing;

the articles in the second listing are associated with the second item by virtue of the

second concept appearing in the articles in the second listing; and

determining whether to compress the second item entry into the first item entry comprises

determining whether the first concept is related to the second concept.

32. The method of claim 31, wherein determining whether the first concept is related to

the second concept comprises accessing a semantic network that stores relationships between

concepts.

33. The article of claim 18, wherein:

the first item comprises a first word;

the articles in the first listing are associated with the first item by virtue of the first word

appearing in the articles in the first listing; and

the first word does not appear in the second item.

34. The article of claim 18, wherein:

the first item comprises a first concept;

the second item comprises a second concept;

the articles in the first listing are associated with the first item by virtue of the first

concept appearing in the articles in the first listing;

the articles in the second listing are associated with the second item by virtue of the

second concept appearing in the articles in the second listing; and

determining whether to compress the second item entry into the first item entry comprises

determining whether the first concept is related to the second concept.

35. The article of claim 34, wherein determining whether the first concept is related to

the second concept comprises accessing a semantic network that stores relationships between

concepts.

Evidence Appendix

None.

Related Proceedings Appendix

None.